219.38418X00
P8768

UNITED STATES PATENT APPLICATION FOR:

METHOD AND APPARATUS FOR PREVENTING PACKET RETRANSMISSIONS DURING IPSEC SECURITY ASSOCIATION ESTABLISHMENT

Inventor:

JAMES L. JASON JR.

Prepared by:

Antonelli, Terry, Stout & Kraus, LLP
1300 North Seventeenth Street, Suite 1800
Arlington, Virginia 22209
Tel: 703/312-6600
Fax: 703/312-6666

METHOD AND APPARATUS FOR PREVENTING PACKET RETRANSMISSIONS DURING IPSEC SECURITY ASSOCIATION ESTABLISHMENT

BACKGROUND

Field

This invention relates to Internet Protocol security (IPsec) protocol secure channels, and more specifically to preventing unnecessary packet retransmissions during IPsec security association establishment.

Background

IPsec is a standard-based network security protocol that is positioned at the network layer (OSI layer 3) of the Transmission Control Protocol/Internet Protocol (TCP/IP) stack. To protect a network flow, IPsec performs processing on outgoing and incoming packets using a security association. This security association describes the network packet flow information (IP addresses, protocol and ports) and the protection suite (algorithms, etc.) used to protect the network packet flow. Variations of IP addresses, protocols, and ports are used to define filters. When a host transmits a packet that matches filter information but is not currently being protected by an active security association, the Internet Key Exchange (IKE) protocol may be used to establish a security association with the communicating peer or network unit. Because of the time necessary for generating keys, network latencies, etc., the IKE negotiation may take some time. Until the security association is finally established, the IPsec packet classification driver has no choice but to discard packets that are to be protected by the security association

under negotiation. This is because of the limited amount of non-paged memory in the operating system (OS) kernel.

For Transmission Control Protocol (TCP) communication, before any application data is sent, the TCP/IP stack sends a sync (SYN) packet. The SYN packet is used to begin the connection establishment procedure. Since TCP is a reliable protocol, when the network layer does not receive an acknowledgment from the communicating peer (because the driver never sent the packet, but instead discarded it), the network layer tries to retransmit the SYN packet. The timeout that TCP uses for retransmitting the packet starts out small. As more retransmissions are required, the timeout is increased.

Fig. 1 shows a diagram of an example system containing a client, a gateway, and a server. Client 10 accesses a gateway 14 using the Internet 12. The gateway 14 protects access to devices on an Intranet 16 including server 18. To send data packets from client 10 to server 18, the client must first send a SYN packet to establish a connection with server 18.

Fig. 2 shows a flow chart of the retransmission process. An application at a host makes a request for a communication to be established to transfer data across a network to another device "communicating peer" (e.g., server) S1. Before the data is sent, a SYN packet is sent to the communicating peer S2. The network layer at the host unit determines if an acknowledgment from the communicating peer has been received, acknowledging that the SYN packet has been received at the communicating peer S3. If an acknowledgment has been received at the host, a communication channel is established and the data is sent S4. However, if an acknowledgment from the communicating peer has not been received, the network layer waits a particular length X of time S5. The network layer then resends the SYN packet S6. Again it is determined if an

acknowledgment has been received from the communicating peer S7, and if so, the data packets are sent S4. If an acknowledgment has not been received from the communicating peer, then the time X is increased S8. The network layer at the host then waits for an amount of time equal to X S9, and then resends the SYN packet again S10. It is then determined whether an acknowledgment was received S11, and if so, the data is sent S4. This process is repeated until a timeout occurs S12, or an acknowledgment is received. If a timeout has occurred, the connection attempt is ended S13.

For an application which uses User Datagram Protocol (UDP) and provides its own reliability, the application will retransmit packets while the driver is dropping the packets and waiting for a security association to be established. Once a security association is established, the TCP/IP stack or the UDP-based application will wait, on the average, half of the current timeout value before sending the packet. As noted previously, trying to address the retransmission problem at the packet classification driver has problems in that because of buffer limitations in the operating system kernel, it is too late to mitigate packet transmissions by the time a packet reaches the IPsec packet classification driver. Further, the classification driver would have store the whole packet, but due to the speed of most systems, this would cause the packet classification driver's memory to fill up quickly causing the host unit to possibly stop.

Therefore, a mechanism is needed that can prevent these unnecessary packet retransmissions.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is further described in the detailed description which follows in reference to the noted plurality of drawings by way of non-limiting examples of embodiments of the present invention in which like reference numerals represent similar parts throughout the several views of the drawings and wherein:

Fig. 1 is a diagram of an example system with a client and server on different networks connected through a gateway;

Fig. 2 is a flowchart of an example retransmission process.

Fig. 3 is a block diagram of components of an example system for preventing packet retransmissions according to an example embodiment of the present invention;

Fig. 4 is a flowchart of an example process for preventing packet retransmissions when an application requests a TCP connection according to example embodiment of the present invention; and

Fig. 5 is a flowchart of an example process for preventing packet retransmissions when an application makes a request for a UDP data transmission according to an example embodiment of the present invention.

DETAILED DESCRIPTION

The particulars shown herein are by way of example and for purposes of illustrative discussion of the embodiments of the present invention. The description taken with the drawings make it apparent to those skilled in the art how the present invention may be embodied in practice.

Further, arrangements may be shown in block diagram form in order to avoid obscuring the invention, and also in view of the fact that specifics with respect to implementation of such block diagram arrangements is highly dependent upon the platform within which the present invention is to be implemented, i.e., specifics should be well within purview of one skilled in the art. Where specific details (e.g., circuits, flowcharts) are set forth in order to describe example embodiments of the invention, it should be apparent to one skilled in the art that the invention can be practiced without these specific details. Finally, it should be apparent that any combination of hard-wired circuitry and software instructions can be used to implement embodiments of the present invention, i.e., the present invention is not limited to any specific combination of hardware circuitry and software instructions.

Although example embodiments of the present invention may be described using an example system block diagram in an example host unit environment, practice of the invention is not limited thereto, i.e., the invention may be able to be practiced with other types of systems, and in other types of environments (e.g., servers).

Reference in the specification to "one embodiment" or "an embodiment" means that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment of the invention. The appearances of the phrase "in one embodiment" in various places in the specification are not necessarily all referring to the same embodiment.

Implementation of a solution to the packet retransmission problem in the TCP/IP stack also has problems. The TCP/IP stack is usually an integral part of the operating system and, therefore, not modifiable unless the company owns the operating system source code. Therefore,

including modifications to correct packet retransmissions would not be possible in the source code.

Moreover, to attempt to correct the packet retransmission problem at the application level, requires that information about the IPsec capabilities be made available to the applications. Also, the applications need to be rewritten to back off of connection establishment or UDP packet transmission until the IPsec security association is established. With this solution, the driver would set up a security association, and then send the data. However, here the advantage of a transparent IPsec implementation would be lost, namely, the fact that IPsec can protect network flows for legacy applications.

In methods and apparatus for preventing packet retransmissions according to the present invention, a network interceptor (i.e., network shim) is placed between the application and the TCP/IP stack. When an application on one unit desires to communicate with another application on another unit across a network, the application uses a socket. A socket is an abstraction that is used to represent one end point of a network communication. Since the network interceptor is between the application and TCP/IP stack, all requests for network communication must go through the network interceptor. The network interceptor can, therefore, monitor specific socket requests to make sure that IPsec security associations are in place before any packets are allowed to flow. Therefore, erroneous packet retransmissions are prevented.

Fig. 3 shows a block diagram of components of an example system according to methods and apparatus for preventing packet retransmissions according to the present invention. All components shown in Fig. 3 reside at a host unit or a network device connected to a network. Application 22 makes requests, for establishment of TCP connection or for transmission of data

on a UDP socket. The application sends its socket request to a network interceptor (i.e., network shim) 24. The network interceptor monitors the application's socket request and may notify a security association negotiation component 30 to negotiate a security association. The network interceptor 24 resides between the application 22 and the TCP/IP stack 26. An IPsec packet classification component 28 interfaces with the TCP/IP stack, and performs the IPsec processing on incoming and outgoing packets. A security association database 32 interfaces with network interceptor 24 and contains a mapping of network flow information (IP addresses, protocol, ports, etc.) to security association information. A security policy database 34 interfaces with network interceptor 24 and contains policies which describe the parameters that are to be used in the negotiation of a security association.

The network interceptor 24 monitors application requests for establishment of a TCP connection, and application requests for transmission of data on a UDP socket. The network interceptor determines if there is a security association already established, and if not, the network interceptor notifies the security association negotiation component, e.g., Internet Key Exchange (IKE), to negotiate one. By monitoring an application's network usage, the network interceptor can insure that security associations are in place before any packets or network traffic begin to flow. Once a security association is established, the application's request is allowed to proceed. For a TCP connection establishment, the TCP/IP stack may be alerted that it needs to establish the connection. In the case of a UDP packet transmission, the packet may be allowed to proceed to the TCP/IP stack which would process the packet for sending.

As noted previously, without the network interceptor, an application creates a TCP socket and does a "connect". This generates a TCP SYN packet that goes to the driver. The driver

realizes that there is no active security association between his machine and the destination machine and, therefore, kicks off an IKE negotiation. During this time, the TCP stack notices that it has not received an acknowledge for the SYN packet, so it sends it again (retransmission).

A network interceptor (i.e., network shim) according to the present invention interrupts the application from performing a "connect". Therefore, the network interceptor blocks the application and tells a negotiation component to perform negotiation for a security association. When a negotiation is completed, the network interceptor lets the "connect" go through to the TCP/IP stack. Now, when the packet hits the driver, a security association is already established. Therefore, instead of the driver dropping the packet (while waiting for a security association to complete), the driver sends the packet out across the network.

When an IP packet comes down the stack to be sent across the network, the packet is compared against filters to determine if a security association is required and needs to be established for protection of the transfer of this packet. A policy is a set of rules. A rule is a condition and its corresponding action. A condition may specify a group of filters which are matched against network traffic (packet) parameters. Filters may be used to match against source address, destination address, source port, destination port, and IP protocol. These filters may be more than just simple values. For example, the source address filter may actually be used to specify an entire subnet. An action is what is to be performed when the corresponding condition evaluates to true (i.e., the network traffic (packet) matches the filters).

When a packet is compared with a filter, it is determined whether a security association is required for protection of the data packet before the data packet is sent across the network.

Thus, the IP packet is compared with all filters at the host unit, and if there is a match, the policy defines the appropriate action to take for that particular filter. For example, if a packet has a destination address and destination port suggesting that it is to be sent to "www.intel.com", when the destination address and destination port information from the packet is compared with a filter, a match may occur whereby the policy defines that 3-DES (Data Encryption Standard) encryption should be suggested during the negotiation of a security association as the encryption technique for transferring the data packet to Intel's server. In contrast, if the data packet has a destination address and destination port suggesting that it is to be sent to "www.cnn.com" this information may be compared against all filters and a match obtained for a particular filter, whereby the policy may define that a match for that particular filter needs no security association established. Therefore, the packet may be transferred across the network to Cable News Network's (CNN) server without any security. Protocol refers to the protocol associated with the data packet that is used on top of the Internet Protocol, for example, Transmission Control Protocol (TCP), User Datagram Protocol (UDP), Internet Control Message Protocol (ICMP), Internet Group Management Protocol (IGMP), etc.

Fig. 4 shows a flow chart of an example process for preventing packet retransmissions when an application requests a TCP connection according to an example embodiment of the present invention. An application makes a request for establishment of a TCP connection S16. It is determined if there is an active security association that covers the network flow associated with the TCP connection request S17. If there is an active security association that covers the network flow, including allowing the communication in the clear, the request is allowed to complete S26, and the TCP connection is established. If there is no active security association that covers the

network flow, it is determined if there is a security policy that requires IPsec for the network flow S18. If there is not a security policy that requires IPsec, a determination is made as to whether the packets associated with the TCP connection request may be allowed to be sent across the network without a security association S19. This may occur if the packet matches a filter in which the corresponding action states to allow the traffic to flow in the clear. If the packets cannot be sent without a security association, the connection request fails S20. If the packets may be allowed to go without a security association, the request is allowed to complete S26.

If there is a security policy that requires IPsec S18, the security negotiation component, i.e., IKE, is notified of the need for establishment of a security association S21. IKE then proceeds with negotiation of a security association S22. It is determined whether IKE (Internet Key Exchange) was successful S23. If IKE was successful, the security association information is saved S25, and the request is allowed to complete S26. If IKE was not successful, the socket is marked accordingly so that negotiation will not be attempted again S24.

Fig. 5 shows a flow chart of an example process for preventing packet retransmissions when an application makes a request for a UDP data transmission according to an example embodiment of the present invention. An application makes a request for a UDP data transmission S30. A determination is made as to whether the socket already has a security association S31. This would apply when there are already transmissions occurring on the requested socket. If the socket already has a security association, the UDP data transmission request is allowed to complete S36. If the socket does not have a security association a determination is made as to whether there is an active security that covers the network flow associated with the UDP data transmission request S32. If there is an active security association

that covers this network flow, the socket is marked and the security association information is saved S35. If there is no active security association that covers the network flow, a determination is made as to whether there is a security policy that requires IPsec for the network flow S33. If there is no security policy that requires IPsec, a determination is made as to whether the UDP data packets may be allowed to be transferred in the clear without a security association S34. If the packets can be transmitted without a security association, the socket is marked and the security association information saved S35. The request for UDP data transmission is then allowed to complete S36. If it is determined that the packets need a security association and, therefore, cannot be allowed to travel without one, the socket is marked such that an attempt to transmit the packet will not occur again S40.

If there is a security policy that requires IPsec, the security negotiation component, i.e., IKE, is notified of the need for establishment of a security association S37. IKE then initiates the negotiation process for establishment of a security association S38. It is determined whether IKE was successful in establishing a security association S39. If IKE was successful, the socket is marked and the security association information is saved S35, and the request is allowed to complete S36. If IKE was unsuccessful in negotiating a security association, the socket is marked such that the process will not be attempted again S40.

Preventing packet retransmissions according to the present invention has several advantages. Packets are no longer silently dropped by the driver causing the TCP/IP stack or the application to generate retransmissions. Further, since the TCP/IP stack and applications use only a finite number of retransmissions, some of these retransmissions are no longer wasted because of the driver silently dropping the packets. Currently, connection establishment may be actually

delayed even more that just the time it takes to establish the security association. TCP uses a retransmission timer approach that employs longer timeout times the more times it has to retransmit. These timeout values become large (respective to the actual network latency) rather quickly. On average, after a security association is established the TCP/IP stack may wait one half of the current timeout value before attempting to send the connection establishment (SYN) packet. According to the present invention, the first TCP SYN packet is not sent until after the security association is in place.

Moreover, the chances of a TCP connection being rejected are reduced. For example, assume that a TCP/IP stack sends N SYN packets before it gives up on trying to establish the connection. Assume that N -1 of those SYN packets are sent, and silently discarded, during the security association establishment. The TCP/IP stack only has one opportunity now to get the SYN packet through. According to the present invention, the TCP/IP stack will still have N attempts to establish the connection. The same applies to a UDP application that provides it own reliability. Since the network interceptor holds onto a send request until after the security association is successful, the application will wait until the network interceptor finishes the send request before starting a timer on the packet.

It is noted that the foregoing examples have been provided merely for the purpose of explanation and are in no way to be construed as limiting of the present invention. While the present invention has been described with reference to a preferred embodiment, it is understood that the words which have been used herein are words of description and illustration, rather than words of limitation. Changes may be made within the purview of the appended claims, as presently stated and as amended, without departing from the scope and spirit of the present

invention in its aspects. Although the present invention has been described herein with reference to particular methods, materials, and embodiments, the present invention is not intended to be limited to the particulars disclosed herein, rather, the present invention extends to all functionally equivalent structures, methods and uses, such as are within the scope of the appended claims.